BAOJ Biotechnology

Nikita Chordia, BAOJ Biotech 2018, 4: 1
4: 031

Review

Reverse Vaccinology: Use of Genomes for Vaccine Design

Nikita Chordia and Anil Kumar*

Abstract

Vaccines have a major impact on public health and life-expectancy. Reverse vaccinology is considered to be fundamental part for vaccine development. Reverse vaccinology includes comparative *in silico* analysis of genome sequences to predict the epitope(s) of pathogen. Further progression includes pan genome, comparative, structural and functional genomics. The process of reverse vaccinology was first applied to serogroup B *Neisseria meningitidis*, thereafter, has been applied to several pathogens including pathogenic *Escherichia coli*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Streptococcus pyogenes*, *Listeria monocytogenes* and has provided scores of new candidate antigens for preclinical and clinical investigation. Reverse vaccinology is fast and efficient but can be applied to only those pathogens whose genome is sequenced. Vaxign, NERVE and Jenner Predict are some of the tools for reverse vaccinology.

Keywords: Vaccinology; Epitope; Pathogen; Antigen; Genome And Proteome.

Introduction

Vaccination is one of the most effective measure for the prevention of infectious diseases. It was started in Asia by Fenner and co workers using materials from smallpox lesions to transmit a mild infection and thereby protect against more serious disease [1]. It was formally introduced by Edward Jenner who used infected materials isolated from cow to immunize against smallpox and he introduced the terminology "vaccine" [2]. The rational development of vaccines was started by Louis Pasteur. He established the basic rules of vaccinology that in order to make a vaccine, one should isolate, inactivate and inject the microorganism that causes the disease [3]. These rules were followed for a century by vaccine developers and many vaccines were developed using this conventional method. This approach is successful in many cases, but is time-consuming and fails when the pathogens cannot be cultivated *in vitro*, or when the most abundant antigens are variable in sequence. Using this technique, many vaccines have been designed by the end of 20th century. Now new technique is required to conquer the remaining pathogens. A revolution came when Craig Venter published the genome of the first free living organism. The availability of complete genome sequences, together with the progression of high-throughput technologies such as functional and structural genomics, has led to a new paradigm in vaccine development. This technology has the capacity to move beyond the rules of Pasteur, using the computer to rationally design vaccines with information present in the genome, without the need to grow the specific microorganisms. This new approach was named as "reverse vaccinology". This was named so as the process of vaccine discovery starts *in silico* using the expressed genomic sequences to find new potential vaccines as compared to other approaches that use the pathogenic organism itself.

Reverse vaccinology defines the process of *in silico* antigen discovery starting from genome information. It was first proposed by Rappuoli in the year, 2000 and represents a genome-based approach to vaccine development [4]. The first successful pathogen addressed using this approach was serogroup b. *meningococcus* (MenB) that causes the meningococcal meningitis worldwide. This bacterium is a challenge for the traditional methods because its capsular polysaccharide is identical to a human self-antigen, whereas the bacterial surface proteins are extremely variable [5]. By the mid 90's many attempts were failed to design vaccine for this bacteria, then availability of the MenB genome sequence gave a hope to design a vaccine [6]. The entire genome sequence of *Neisseria meningitidis* was computer-analyzed and *in silico* selected 600 novel vaccine candidates were expressed in *E. coli*. Out of them, 350 were successfully expressed and purified. These were used for immunization of mice and testing of sera for bactericidal activity *in vitro* (complement-mediated *in vitro* killing of the bacteria). Finally, 29 novel, surface-exposed proteins were shown to induce bactericidal antibodies and most of these vaccine

***Corresponding Author:** Anil Kumar, School of Biotechnology, Devi Ahilya University, Khandwa Rd, Indore-452001, India, E-mail: ak_sbt@yahoo.com

Sub Date: June 13th, 2018, **Acc Date:** June 22nd, 2018, **Pub Date:** June 26th, 2018.

Citation: Nikita Chordia and Anil Kumar (2018) Reverse Vaccinology: Use of Genomes for Vaccine Design. BAOJ Biotech 4: 031.

candidates were not discovered by all previously used techniques [7]. The antigens inducing the best and broadest bactericidal activity were selected and inserted into prototype vaccines that were able to induce protective immunity against most of the MenB strains in mice [8]. After successful preclinical studies, the MenB vaccine entered the long path of vaccine development. Since then, it has been used for several other pathogens. In this review, we have summarized the approach, refinements, progress of reverse vaccinology with its advantages, disadvantages and lessons. We have also discussed the available tools for the reverse vaccinology.

Approach

Reverse vaccinology or sometimes called epitope-driven vaccine design uses genomic information to derive vaccine candidates from *in silico* analysis. Genomic databases generally contain the whole genome sequences with the complete repertoire of proteins encoded by the genomes which have made it possible to screen vaccine candidates. The process starts with the gene finding to distinguish coding from non-coding microbial DNA. Another important criterion is to perform inter species comparative genomics analyses to ensure that there is no cross-reacting between pathogenic antigens and human proteins. To increase vaccine specificity, the analysis can be refined further by excluding antigens present in other pathogens or any related bacterial strains. Proteins selected by genome-based approach are further analyzed for finding surface-exposed or secreted antigens. These are good vaccine candidates due to their increased accessibility to the immune system. This is done by deducing their putative sub-cellular localization spanning from cytoplasm to the cell wall (Gram positive bacteria) or outer membrane (Gram negative bacteria) according to their specific signatures like trans-membrane helices, secretory leader sequences, cell-wall anchor motifs etc [9]. For the selected protein, B and T cell epitopes are identified. The B cells are important in recognizing the epitopes of the antigen that can be identified by the paratopes of antibody. In some cases, T cells play a role in cell mediated immunity as the processed antigenic peptides interact with the T cell when they are presented in context of T cell. The epitope prediction plays an important role in designing of epitope based vaccine [10].

Refinements in Reverse Vaccinology

Pan Genomic Reverse Vaccinology

Pan Genome refers to the sum of the core genome shared by all sequenced strains and the dispensable genome present only in a subset of the isolates. It is analyzed to assess the size and diversity of the gene repertoire that the species has access to [11]. In Pan genomic reverse vaccinology, the genomes of the different isolates of same organism are compared with each other by using computer analysis. The first pan genome approach was done against *Streptococcus agalactiae* [12].

Comparative Reverse Vaccinology

In this approach, pathogenic and non pathogenic strains of one species are compared at their genetic level. It deals with the differences in structure of proteins of different organisms.

Applications

Meningitis

As described above, the development of a serogroup b *Neisseria meningitides* (MenB) vaccine represents the first example of the successful application of reverse vaccinology. Meningitis is the swelling of the membranes around the spinal cord and brain showing the symptoms like fever, stiff neck and back, confusion and coma. After the failure of classical method, Rappuoli applied reverse vaccinology for removing bacterial meningitis in 2001 [13]. Later in 2006, Giuliani and co workers designed universal vaccine for meningitis. In this work, most of the antigens selected *in silico* were successfully expressed in *Escherichia coli*. A list of five selected antigens of this study were tested with the adjuvant achieving antibodies against more than 90% of 85 strains of *Meningococci* representative of the global population diversity [8].

Listerosis

Listerosis is an infectious food borne disease caused by *Listeria monocytogenes* in animals and human. Its infection can lead to septicemia, meningitis, encephalitis, corneal ulcer, pneumonia, and intrauterine or cervical infections. Reverse vaccinology approach has been used for the development of the vaccine against listeriosis [14]. They used many programs like Signal P 3.0, LipoP1.0, TMHMM, PSORTB, HLA Pred and BLASTP in together to search out highly conserved surface antigens as lipoprotein and cell wall anchored proteins and highlighted 104 prominent surface antigens which may be involved in subunit vaccine development programs.

Rickettsiosis

Rickettsiosis is a group of diseases caused by many intracellular bacteria of the genus *Rickettsia*. The infection in human occurs either from a tick bite or rarely by contamination of cut skin or a wound with faeces of the ticks. Common vaccine candidates for the six rickettsial species were designed by Chordia et al. in 2016 [15]. A total 19 proteins were selected which were predicted to be localized on the membrane having less than or one transmembrane helices. On further analysis, total 10 potential epitopes were identified as potent vaccine candidates for rickettsiosis.

Pneumoniae

Streptococcus pneumoniae is a gram-positive bacterium that causes significant human disease, including sepsis, meningitis, pneumonia, otitis media and sinusitis. It accounts for 11% of mortality worldwide in children under 5 years of age. The availability of multiple pneumococcal

genome sequences combined with the increased understanding of pili in *Streptococcus agalactiae* and *Streptococcus pyogenes* led to the discovery of pili in *S. pneumoniae* [16,17]. The investigation led to find the pneumococcal pilus proteins as vaccine antigens [18].

Chancroid

It is a sexually transmitted infection caused by the Gram-negative bacterium *Haemophilus ducreyi*. Vaccine targets are searched using reverse vaccinology with subtractive genomics approach to record the cell surface antigens and their epitopes for the high scored values as per conserved nature and ability to span plasma membrane and cell wall. Potential vaccine and drug targets against 28 strains of *H. ducreyi* were searched, 847 non-host homologous proteins, being 332 exposed/secreted/membrane and 515 cytoplasmic proteins were identified. On checking their essentiality, functionality and virulence, 13 candidate vaccine targets and three drug targets were identified [19].

Dengue

Dengue fever caused by the infection of any of the four serotypes of dengue virus (DEN1, DEN 2, DEN 3 and DEN 4). Vaccines available for dengue are specific for particular serotypes. A common vaccine for all four serotypes is designed using reverse vaccinology. All pathogenic proteins from four strains were predicted for better antigenicity. On comparing antigenic regions binding affinities with maximum number of MHC I and II alleles, antigenic peptides were selected [20].

Leishmaniasis

Leishmaniasis is a group of diseases with a spectrum of clinical manifestations ranging from cutaneous ulcers to visceral leishmaniasis, which results from the bite of an infected sandfly to human. Reverse Vaccinology approach is used to identify common vaccine candidates from both highly pathogenic species *Leishmania major* and *Leishmania infantum*. Total proteomes of both species were used for analysis and 19 potential epitopes were found to be potential candidates for vaccine, which can be further verified through in vivo experiments in MHC compatible animal models [21].

Cystoisosporosis

Cystoisospora suis is a coccidian species that typically affects suckling piglets. Infections occur by oral uptake of oocysts and are characterized by non-hemorrhagic transient diarrhea, resulting in poor weight gain. Using Next Generation Sequencing ~84Mb sequence assembly for the *Cystoisospora suis* genome is produced and 1,168 vaccine candidates are identified in the predicted *Cystoisospora suis* proteome. Further candidates are characterized according to function, conservation, expression and overlap with candidates that had been tested in other coccidians. Only 22% of the candidates with orthologs in *Cystoisospora suis* had a high score and finally by overlapping the vaccine candidates, a promising new vaccine candidate corresponding to a 42 kDa transmembrane protein with unknown function [22].

Reverse vaccinology has been applied to a wide range of bacterial pathogens and has provided a long list of promising antigens from functionally blind interrogation of their genomes, and the subsequent studies on antigen function are leading to increased understanding of the biology of the pathogens. Delany et al. have provided a list of pathogens with their vaccine development [23].

Advantages

The major advantage for reverse vaccinology is that it is very quick and efficiently finds vaccine targets. Traditional methods may take decades to unravel diseases, pathogens, antigens and immunity. However, in silico approach can be very fast, allowing to identify new vaccines for testing in only a few years. Reverse vaccinology uses the entire protein sequences of each pathogen to select the best candidate vaccine antigens rather than only an unsuitable subset expressed by a cell line. This approach allows the development of vaccines that were previously difficult or impossible to make and can lead to the discovery of unique antigens that may improve existing vaccines. The conventional way of vaccine development includes culturing of pathogens in laboratory but this is not possible in case of highly infectious pathogens that are hazardous to culture in laboratory. This problem is resolved by in silico method to design an efficient vaccine. This is a low cost technique, fully feasible of use into the plethora of genomic data being generated [24].

Disadvantages

Despite the extensive use of the initial concept of reverse vaccinology and advanced rated results in the search for vaccines against certain pathogens, in general its best and most practical results are still expected. Using this method, only proteins can be targeted. Whereas, conventional vaccinology approaches can find other bimolecular targets such as polysaccharides. One of the major downside of this method is that it can be applied only to those pathogens whose genome sequence is available. Despite being successful for bacterial vaccines, it remains an overseen approach for other classes of pathogens such as protozoa, helminthes and viruses [25].

Lessons from Reverse Vaccinology

The basic lessons from this approach are identification of candidate antigens on a pure bioinformatics base irrespective of their functions are still alive and generally applicable. Downstream from this step, advancements in technologies, especially the omics approaches, have improved the efficiency of antigen identification, selection and engineering, allowing the use of experimental data to complement the bioinformatics predictions. Still, comprehension of the mechanisms of infection of many pathogens is crucial in order to develop correlates of protection and predict not only immunogenicity but also protective efficacy and the quality of the immune response. The next lessons could come from the discovery of the expanded coding potential of the genomes throughout non-canonical ORFs and the potential relevance of those proteins for an efficacious human immune response [26].

Software For Reverse Vaccinology

Vaxign

Vaxign is a vaccine target prediction and analysis system based on the principle of reverse vaccinology. It is a freely available program available at http://www.violinet.org/vaxign/ that facilitates vaccine researchers to efficiently design vaccine targets. It predicts possible vaccine targets based on various vaccine design criteria using microbial genomic and protein sequences as input data. Major predicted features include transmembrane domain, sequence conservation among genomes, adhesion probability, sequence similarity to host (human or mouse) proteome, sub-cellular localization of a protein, and epitope binding to MHC class I and class II. The following two programs exist in Vaxign:

Vaxign Query: This program allows users to search precomputed Vaxign results. The precomputed Vaxign database contains prediction of vaccine targets for more than 70 genomes.

Dynamic Vaxign *Analysis*: This program allows users to input a protein sequence(s) and set up parameters. The Vaxign web server dynamically calculates the possibilities of using the protein(s) as vaccine target(s) [27].

Jenner Predict

Jenner-Predict server is freely accessible at http://117.211.115.67/vaccine/home.html. It has been developed for prediction of protein vaccine candidates (PVCs) from proteomes of bacterial pathogens. The web server targets host-pathogen interactions and pathogenesis by considering known functional domains from protein classes such as virulence, invasin, adhesin, porin, flagellin, toxin, choline-binding, colonization, penicillin-binding, fibronectin-binding, transferring-binding and solute-binding. It predicts non-cytosolic proteins containing above domains as PVCs. It also provides vaccine potential of PVCs in terms of their possible immunogenicity by comparing with experimentally known immune epitope database (IEDB) epitopes, absence of autoimmunity and conservation in different strains. Predicted PVCs are prioritized so that only few prospective PVCs could be validated experimentally [28].

Nerve

New Enhanced Reverse Vaccinology Environment (NERVE) is user-friendly software available at http://www.bio.unipd.it/molbinfo for the *in silico* identification of the best vaccine candidates from whole proteomes of bacterial pathogens. The software integrates multiple robust and well-known algorithms for protein analysis and comparison. Vaccine candidates are ranked and presented in a HTML table showing relevant information and links to corresponding primary data. NERVE has been tuned as the first available tool able to rank a restricted pool (~8–9% of the whole proteome) of vaccine candidates and to show high recall (~75–80%) of known protective antigens. These vaccine candidates are required to be "safe" (taking into account autoimmunity risk) and "easy" for further experimental, high-throughput screening (avoiding possibly not soluble antigens) [29].

Conclusion

The whole genome sequence is required for the prediction of epitopes and other surface proteins which is the important part of reverse vaccinology for the designing of a successful candidate vaccine. The examples given in this review indicated that reverse vaccinology is useful for the preparation of epitope based vaccines against the most dangerous pathogens. Reverse vaccinology predicts the vaccine candidate for several pathogens but still promising vaccine against some pathogens seems to be far away because genome sequences of most of the major human pathogens are not available.

Acknowledgements

The facilities of the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi (DBT) under the Bioinformatics Sub Centre used in the present work are gratefully acknowledged.

References

1. Fenner F, Henderson DA, Arita I, Jezek Z, Ladnyi ID & World Health Organization (1988) Smallpox and its eradication.
2. Jenner E (1801) The Origin of the Vaccines Inoculation (London: Shury)
3. Pasteur L (1880) De l'attenuation du virus du Chole´ ra des poules. CR Acad Sci Paris 91: 673–680
4. Rappuoli R (2000) Reverse vaccinology. Curr Opin Microbiol 3: 445–450
5. Sette A, Rappuoli R (2010) Reverse vaccinology: developing vaccines in the era of genomics. Immunity 33(4): 530-541.
6. Tettelin H, Saunders NJ, Heidelberg J, Jeffries AC, Nelson KE, et al. (2000) Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. Science 287(5459): 1809–1815.
7. Pizza M, Scarlato V, Masignani V, Giuliani MM, Aricò B, et al. (2000) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 287(5459): 1816-1820.
8. Giuliani MM, Adu-Bobie J, Comanducci M, Arico` B, Savino S, et al. (2006) A universal vaccine for serogroup B meningococcus. Proc. Natl. Acad. Sci. USA 103(29): 10834–10839.
9. Bertholet S, Reed SG, Rappuoli R (2014) Reverse vaccinology applied to TB. The art & science of tuberculosis vaccine development: 413-431.
10. Kanampalliwar AM, Soni R, Girdhar A, Tiwari A (2013) Reverse vaccinology: basics and applications. J Vaccines Vaccin 4(6): 194-198.

11. Tettelin H (2009) The bacterial pan-genome and reverse vaccinology. In Microbial Pathogenomics 6:35-47
12. Lefébure T, Stanhope MJ (2007) Evolution of the core and pan-genome of Streptococcus: positive selection, recombination, and genome composition. Genome biology 8(5): R71.
13. Rappuoli R (2001) Conjugates and reverse vaccinology to eliminate bacterial meningitis. Vaccine 19(17-19): 2319-2322.
14. Gore D, Pachkawade M (2012) In silico Reverse Vaccinology Approach for Vaccine Lead Search in Listeria monocytogenes. Biocompx 1: 15-22.
15. Chordia N, Choudhary S, Kumar A (2016) Identifcaton of Potential Vaccine Candidates from Ricketssia Species: A Reverse Vaccinology Approach. BAOJ Biotech 2(1): 006.
16. Barocchi MA, Ries J, Zogaj X, Hemsley C, Albiger B, et al. (2006) A pneumococcal pilus influences virulence and host inflammatory responses. Proc Natl Acad Sci USA 103(8): 2857–2862.
17. LeMieux J, Hava DL, Basset A, Camilli A (2006) RrgA and RrgB are components of a multisubunit pilus encoded by the Streptococcus pneumonia rlrA pathogenicity islet. Infect Immun 74(4): 2453–2456.
18. Talukdar S, Zutshi S, Prashanth KS, Saikia KK, Kumar P (2014) Identification of potential vaccine candidates against Streptococcus pneumoniae by reverse vaccinology approach. App biochemi biotech 172(6): 3026-3041.
19. de Sarom A, Jaiswal AK , Tiwari S, de Castro Oliveira L, Barh D, et al. (2018) Putative vaccine candidates and drug targets identified by reverse vaccinology and subtractive genomics approaches to control Haemophilus ducreyi, the causative agent of chancroid. J R Soc Interface 15(142): 20180032.
20. Baskar V , Madhan R, Srinivasan G, Selvakumar K, Radha M (2011) Identifying the Potential Tetravalent Vaccine Candidate for Dengue Virus using Insilico Approach. Insight Bioinfo. 1(1): 6-10.
21. John L, John GJ, Kholia T (2012) A reverse vaccinology approach for the identification of potential vaccine candidates from Leishmania spp. App biochem biotech 167(5): 1340-1350.
22. Palmieri N, Shrestha A, Ruttkowski B, Beck T, Vogl C, et al. (2017) The genome of the protozoan parasite Cystoisospora suis and a reverse vaccinology approach to identify vaccine candidates. Int J Parasitology 47(4): 189-202.
23. Delany I, Rappuoli R, Seib KL (2013) Vaccines, reverse vaccinology, and bacterial pathogenesis. Cold Spring Harb Perspect Med 3(5): a012476.
24. Santos A, Ali A, Barbosa E, Silva A, Miyoshi A, et al. (2011) The reverse vaccinology-A contextual overview. IIOABJ 2: 8-15.
25. Del Tordello E, Rappuoli R, Delany I (2017) Reverse Vaccinology: Exploiting Genomes for Vaccine Design. In Human Vaccines 65-86.
26. Bruno L, Cortese M, Rappuoli R, Merola M (2015) Lessons from Reverse Vaccinology for viral vaccine design. Curr Opin virology 11: 89-97.
27. Xiang Z, He Y (2009) Vaxign: a web-based vaccine target design program for reverse vaccinology. Procedia in Vaccinology 1(1): 23-29.
28. Jaiswal V, Chanumolu SK, Gupta A, Chauhan RS, Rout C (2013) Jenner-predict server: prediction of protein vaccine candidates (PVCs) in bacteria based on host-pathogen interactions. BMC bioinformatics 14(1): 211.
29. Vivona S, Bernante F, Filippini F (2006) NERVE: new enhanced reverse vaccinology environment. BMC biotechnology 6(1): 35.